

Homework Assignment 5: Bayes Networks

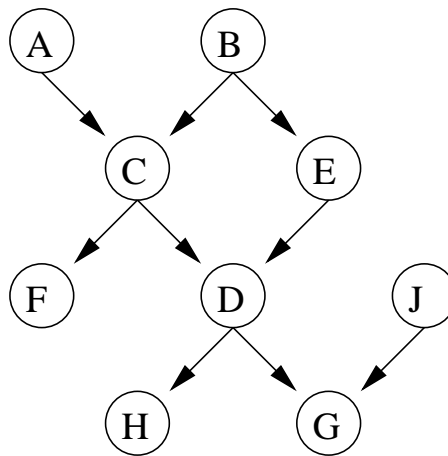
CSC 384 – Winter 2003

Out: March 26, 2003

Due: April 7, 2003: in class

Be sure to include your name and student number with your assignment.

1. [45 pts] Consider the following Bayes net (we leave out the Conditional Probability Tables).



For each of the questions below, state whether the independence relation holds. Give a justification for your answer using the d-separation criterion: if d-separation holds, state why each *undirected* path is blocked; if it does not hold, describe one undirected path which is unblocked. (Note: there are no more than two undirected paths between any pair of nodes).

- (a) [3 pts] A and B are independent
- (b) [3 pts] A and B are independent given D
- (c) [3 pts] A and E are independent
- (d) [3 pts] A and E are independent given C
- (e) [3 pts] A and E are independent given B
- (f) [3 pts] A and E are independent given B and C
- (g) [3 pts] A and E are independent given D
- (h) [3 pts] A and E are independent given F
- (i) [3 pts] F and H are independent
- (j) [3 pts] J and E are independent
- (k) [3 pts] J and E are independent given G
- (l) [3 pts] J and E are independent given A
- (m) [3 pts] G and A are independent
- (n) [3 pts] G and A are independent given C
- (o) [3 pts] G and A are independent given C and D

2. **[55 pts]** Every year, credit card companies lose millions of dollars due to frauds resulting from lost or stolen cards. Recently, the financial industry has turned to AI for solutions to the fraud detection problem. Intuitively, credit card holders tend to make purchases following a certain pattern. A fraud is likely to happen when this pattern is broken.

Suppose you are working for a financial institution and you are asked to implement a fraud detection system. You plan to use the following information:

- When the card holder is travelling abroad, fraudulent transactions are more likely since tourists are prime targets for thieves. More precisely, 0.5% of transactions are fraudulent when the card holder is travelling, where as only 0.1% of the transactions are fraudulent when she is not travelling. On average, 5% of all transactions happen while the card holder is travelling. If a transaction is fraudulent, then the likelihood of a foreign purchase increases, unless the card holder happens to be travelling. More precisely, when the card holder is not travelling, 10% of the fraudulent transactions are foreign purchases where as only 1% of the legitimate transactions are foreign purchases. On the other hand, when the card holder is travelling, then 90% of the transactions are foreign purchases regardless of the legitimacy of the transactions.
- Purchases made over the internet are more likely to be fraudulent. This is especially true for card holders who don't own any computer. Currently, 60% of the population owns a computer and for those card holders, 1% of their legitimate transactions are done over the internet, however this percentage increases to 2% for fraudulent transactions. For those who don't own any computer, a mere 0.1% of their transactions is done over the internet, but that number increases to 1.1% for fraudulent transactions. Unfortunately, the credit card company doesn't know whether a card holder owns a computer, however it can usually guess by verifying whether any of the recent transactions involve the purchase of computer related accessories. In any given week, 10% of those who own a computer purchase with their credit card at least one computer related item as opposed to just 0.1% of those who don't own any computer.

- (a) **[25 pts]** Construct a Bayes Network that your fraud detection system can use to identify fraudulent transactions. In particular, you are to show the graph defining the network and the Conditional Probability Tables associated with each node in the graph. This network should encode the information stated above. Your network should contain exactly six nodes, corresponding to the following binary random variables:

- *OC* – card holder owns a computer.
- *Fraud* – current transaction is fraudulent.
- *Trav* – card holder is currently travelling.
- *FP* – current transaction is a foreign purchase.
- *IP* – current purchase is an internet purchase.
- *CRP* – a computer related purchase was made in the past week.

The arcs defining your Bayes Network should accurately capture the probabilistic dependencies between these variables.

- (b) **[12 pts]** What is the prior probability (i.e., before we search for previous computer related purchases and before we verify whether it is a foreign and/or an internet purchase) that the current transaction is a fraud? What is the probability that the current transaction is a fraud once we have verified that it is a foreign transaction, but not an internet purchase and that the card holder purchased computer related accessories in the past week? Show your computation steps. *You may wish to run variable elimination to determine these probabilities.*
- (c) **[8 pts]** After computing those probabilities, the fraud detection system raises a flag and recommends that the card holder be called to confirm the transaction. An agent calls at the domicile of the card holder but she is not home. Her spouse confirms that she is currently out of town on a business trip. How does the probability of a fraud change based on this new piece of information? Show your computation steps.
- (d) **[10 pts]** Suppose you are not a very honest employee and you just stole a credit card. You know that the fraud detection system uses the Bayes net designed earlier but you still want to make an important purchase over the internet. What can you do prior to your internet purchase to reduce the risk that the transaction will be rejected as a possible fraud? Explain.